

Normalized neural representations of natural odors

David Zwicker^{1,2,*}

¹*School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA*

²*Kavli Institute for Bionano Science and Technology, Harvard University, Cambridge, MA 02138, USA*

(Dated: August 4, 2016)

The olfactory system removes correlations in natural odors using a network of inhibitory neurons in the olfactory bulb. It has been proposed that this network integrates the response from all olfactory receptors and inhibits them equally. However, how such global inhibition influences the neural representations of odors is unclear. Here, we study a simple statistical model of this situation, which leads to concentration-invariant, sparse representations of the odor composition. We show that the inhibition strength can be tuned to obtain sparse representations that are still useful to discriminate odors that vary in relative concentration, size, and composition. The model reveals two generic consequences of global inhibition: (i) odors with many molecular species are more difficult to discriminate and (ii) receptor arrays with heterogeneous sensitivities perform badly. Our work can thus help to understand how global inhibition shapes normalized odor representations for further processing in the brain.

I. INTRODUCTION

Sensory systems encode information efficiently by removing redundancies present in natural stimuli (Barlow, 2001, 1961). In natural images, for instance, neighboring regions are likely of similar brightness and the image can thus be characterized by the regions of brightness changes (Ruderman and Bialek, 1994). This structure is exploited by ganglion cells in the retina that respond to brightness gradients by receiving excitatory input from photo receptors in one location and inhibitory input from the surrounding (Demb and Singer, 2015). This typical center-surround inhibition results in neural patterns that represent natural images efficiently (Carandini and Heeger, 2012). Similarly, such local inhibition helps separating sound frequencies in the ear and locations touched on the skin (Isaacson and Scanziani, 2011). Vision, hearing, and touch have in common that their stimulus spaces have a metric for which typical correlations in natural stimuli are local. Consequently, local inhibition can be used to remove these correlations and reduce the high-dimensional input to a lower-dimensional representation.

The olfactory stimulus space is also high-dimensional, since odors are comprised of many molecules at different concentrations. Moreover, the concentrations are also often correlated, e.g., because the molecules originate from the same source. However, these correlations are not represented by neighboring neurons in the olfactory system, since there is no obvious similarity metric for molecules that could be used to achieve such an arrangement (Nikolova and Jaworska, 2003). Because the olfactory space lacks such a metric, local inhibition cannot be used to remove correlations to form an efficient representation (Murthy, 2011; Soucy *et al.*, 2009). Consequently,

the experimentally discovered inhibition in the olfactory system (Yokoi, Mori, and Nakanishi, 1995) likely affects neurons irrespective of their location. Such global inhibition could for instance normalize the activities by their sum, which has been observed experimentally (Olsen, Bhandawat, and Wilson, 2010; Roland *et al.*, 2016). This normalization cannot reduce the correlation structure of odors, but it could help separating the odor composition (what is present?) from the odor intensity (how much is there?) (Cleland, 2010; Laurent, 1999). This separation is useful, since the composition identifies an odor source, while the intensity information is necessary for finding or avoiding it. However, how global inhibition shapes such a bipartite representation of natural odors is little understood.

In this paper, we study a simple model of the olfactory system that resembles its first processing layers, which transform the odor representation successively (Silva Teixeira, Cerqueira, and Silva Ferreira, 2016; Wilson, 2013), see Fig. 1. Our model connects previous results from simulations of the neural circuits (Cleland and Sethupathy, 2006; Getz and Lutz, 1999; Li, 1990, 1994; Linster and Hasselmo, 1997; Zhang, Li, and Wu, 2013) to system-level descriptions of the olfactory system (Hopfield, 1999; Koulakov, Gelperin, and Rinberg, 2007; Zwicker, Murugan, and Brenner, 2016). The main feature of the model is global inhibition, which leads to normalization. This separates the odor composition from its intensity and encodes it in a sparse representation. The inhibition strength controls the trade-off between the sparsity and the transmitted information, which influences how well this code can be used to discriminate odors in typical olfactory tasks. The model reveals two generic consequences of global inhibition: (i) odors comprised of many different molecules exhibit sparser representations and should thus be more difficult to distinguish and (ii) overly sensitive receptors could dominate

* <http://www.david-zwicker.de>

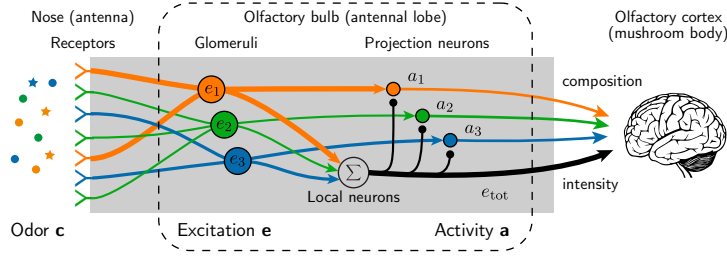


FIG. 1 Schematic picture of our model describing the signal processing in the olfactory bulb: An odor comprised of many ligands excites the olfactory receptors and the signals from all receptors of the same type are accumulated in respective glomeruli. Associated projections neurons receive excitatory input from a single glomerulus and are subject to global inhibition, mediated by a network of local neurons. The activity of the projection neurons form a sparse, concentration-invariant odor representation.

the sparse responses and arrays with heterogeneous receptors should thus perform poorly.

II. SIMPLE MODEL OF THE OLFACTORY SYSTEM

Odors are blends of odorant molecules that are ligands of the olfactory receptors. We describe an odor by a vector $\mathbf{c} = (c_1, c_2, \dots, c_{N_L})$ that specifies the concentrations c_i of all N_L detectable ligands ($c_i \geq 0$). Generally, only a small subset of the $N_L \sim 10^5$ ligands are present in natural odors, so most of the c_i will typically be zero. The ligands in an odor are detected by olfactory receptor neurons, which reside in the nose in mammals and in the antenna in insects (Kaupp, 2010). Each of these neurons expresses receptors of one of N_R genetically defined types, where $N_R \approx 50$ for flies (Wilson, 2013), $N_R \approx 300$ for humans (Verbeurg *et al.*, 2014), and $N_R \approx 1000$ for mice (Niimura, 2012). The excitation of all receptor neurons of the same type is accumulated in associated glomeruli (Su, Menuz, and Carlson, 2009), whose excitation pattern forms the first odor representation, see Fig. 1. Here, the large number of ligands and their possible mixtures are represented by a combinatorial code, where each ligand typically excites multiple receptor types (Malnic *et al.*, 1999). It has been shown experimentally that the excitation e_n of the glomerulus associated with receptor type n can be approximated by a linear function of the ligand concentrations \mathbf{c} (Gupta, Albeanu, and Bhalla, 2015; Silbering and Galizia, 2007; Tabor *et al.*, 2004),

$$e_n = \sum_{i=1}^{N_L} S_{ni} c_i, \quad (1)$$

where S_{ni} denotes the sensitivity of glomerulus n to ligand i . We here consider a statistical description of combinatorial coding by studying random sensitivity matrices with entries drawn independently from a log-normal distribution. This distribution is parameterized by the mean sensitivity \bar{S} and the standard deviation λ of the

underlying normal distribution. This choice is motivated by experimental measurements, which also suggest that $\lambda \approx 1$ for flies and humans (Zwicker, Murugan, and Brenner, 2016). We showed previously that such random matrices typically decorrelate stimuli and thus lead to near-optimal odor representations on the level of glomeruli (Zwicker, Murugan, and Brenner, 2016).

In contrast to our previous model, we here consider the odor representation encoded by projection neurons (mitral and tufted cells in mammals), which constitute the next layer after the glomeruli, see Fig. 1. Projection neurons typically receive excitatory input from a single glomerulus (Jefferis *et al.*, 2001) and inhibitory input from many local neurons (granule cells in mammals), which are connected to other projection neurons and glomeruli (Cleland, 2010; Su, Menuz, and Carlson, 2009). The activity a_n of the projection neurons associated with receptor type n is a sigmoidal function of ligand concentrations (Bhandawat *et al.*, 2007; Tan *et al.*, 2010). Additionally, all signals are subject to noise, both from stochastic ligand-receptor interactions and from internal processing (Lowe and Gold, 1995), which limits the number of distinguishable output activities. We capture both effects by considering the simple case where only two activities a_n can be distinguished. Here, the projection neurons are active when their excitatory input, the respective excitation e_n , exceeds a threshold γ ,

$$a_n = \begin{cases} 0 & e_n \leq \gamma \\ 1 & e_n > \gamma \end{cases}. \quad (2)$$

Generally, γ could depend on the type n , but we here consider a simple mean-field model, where all types exhibit the same threshold. Nevertheless, this threshold could still depend on global variables. Experimental data (Asahina *et al.*, 2009; Aungst *et al.*, 2003; Banerjee *et al.*, 2015; Berck *et al.*, 2016; Hong and Wilson, 2015; Olsen, Bhandawat, and Wilson, 2010; Roland *et al.*, 2016; Silbering and Galizia, 2007) and modeling of the local neurons (Cleland, 2010; Cleland and Sethupathy, 2006) suggest that the total excitation of all glomeruli inhibits all projection neurons. To capture this we postulate that

the threshold γ is a function of the total excitation, where we for simplicity consider a linear dependence,

$$\gamma = \frac{\alpha}{N_R} \sum_{n=1}^{N_R} e_n. \quad (3)$$

Here, α is a parameter that controls the inhibition strength.

Taken together, our model of the olfactory system comprises N_R communication channels, each consisting of receptors, a glomerulus, and projection neurons, which interact via global inhibition, see Fig. 1. The Eqs. 1–3 describe how this system maps an odor \mathbf{c} to an activity pattern $\mathbf{a} = (a_1, a_2, \dots, a_{N_R})$. The amount of information that can be learned about \mathbf{c} by observing \mathbf{a} is quantified by the mutual information I , which reads

$$I = - \sum_{\mathbf{a}} P(\mathbf{a}) \log_2 P(\mathbf{a}). \quad (4)$$

Here, the probability $P(\mathbf{a})$ of observing output \mathbf{a} is given by $P(\mathbf{a}) = \int P(\mathbf{a}|\mathbf{c})P_{\text{env}}(\mathbf{c})d\mathbf{c}$. The conditional probability $P(\mathbf{a}|\mathbf{c})$ of observing \mathbf{a} given \mathbf{c} describes the processing in the olfactory system and follows from the Eqs. 1–3. In contrast, $P_{\text{env}}(\mathbf{c})$ denotes the probability of encountering an odor \mathbf{c} , which depends on the environment. Consequently, the information I is not only a function of the sensitivity matrix S_{ni} and the inhibition strength α , but also of the environment in which the receptors are used (Zwicker, Murugan, and Brenner, 2016).

Natural odor statistics are hard to measure (Wright and Thomson, 2005) and we thus cannot infer the distribution $P_{\text{env}}(\mathbf{c})$ from experimental data. Instead, we consider a broad class of distributions parameterized by a few parameters. For simplicity, we only consider uncorrelated odors, where the concentrations c_i of ligands are independent. We denote by p_i the probability that ligand i is part of an odor. If this is the case, the associated c_i is drawn from a log-normal distribution with mean μ_i and standard deviation σ_i . This choice allows us to independently adjust the mean odor size $s = \sum_i p_i$, the mean of the total concentration $c_{\text{tot}} = \sum_i c_i$, and the concentration variations $\frac{\sigma_i}{\mu_i}$. Averaged over all odors, c_i then has mean $\langle c_i \rangle = p_i \mu_i$ and variance $\text{var}(c_i) = (p_i - p_i^2)\mu_i^2 + p_i \sigma_i^2$. Note that typical odors can have hundreds of different ligands (Wright and Thomson, 2005), but this is still well below $N_L \sim 10^5$ and we thus have $1 \ll s \ll N_L$.

III. RESULTS

A. Global inhibition leads to concentration-invariant, sparse representations

Our model has the interesting property that the odor representation \mathbf{a} does not change when the odor \mathbf{c} or the sensitivities S_{ni} are scaled by a positive factor. This

is because both the excitations e_n and the threshold γ are linear in \mathbf{c} and S_{ni} , see Eqs. 1 and 3, and the activities a_n only depend on the ratio e_n/γ , see Eq. 2. In fact, these equations can be interpreted as normalization of the excitations by the total excitation followed by thresholding with the constant threshold α/N_R . Since the representation \mathbf{a} does not depend on c_{tot} , it only encodes relative ligand concentrations, i.e., the odor composition. This property is called concentration invariance and corresponds to the everyday experiences that odors smell the same over many orders of magnitude in concentration (Cleland *et al.*, 2011; Uchida and Mainen, 2007; Zhang, Li, and Wu, 2013). Indeed, experiments suggest that the activity of projection neurons is concentration-invariant (Cleland *et al.*, 2007; Sachse and Galizia, 2002; Sirotin, Shusterman, and Rinberg, 2015) and exhibits more uniform distances between odors (Bhandawat *et al.*, 2007; Cleland *et al.*, 2007), indicating that they encode the odor composition efficiently.

To understand how odor compositions are encoded in our model, we start with numerical simulations of Eqs. 1–3 as described in the SI. Fig. 2A shows the excitations e_n corresponding to an arbitrary odor. Here, the excitation threshold is 1.4 times the mean excitation, and only three channels are active (orange bars). The corresponding histogram in Fig. 2B shows that the number of active channels is typically small for this inhibition strength when odors are presented with statistics $P_{\text{env}}(\mathbf{c})$. Moreover, the magnitude of the Pearson correlation coefficient between two channels is typically only 1%, see SI. This weak correlation is expected for the uncorrelated odors and random sensitivity matrices that we consider here and explains why the histogram in Fig. 2B is close to a binomial distribution. The odor representations are thus mainly characterized by the mean channel activity $\langle a_n \rangle$.

The mean channel activity $\langle a_n \rangle$ depends on the inhibition strength α , the sensitivities S_{ni} , and the odor statistics $P_{\text{env}}(\mathbf{c})$. To discuss these dependences, we next introduce an approximation based on a statistical description of the associated excitation e_n . Here, we define the normalized concentrations $\hat{c}_i = c_i/c_{\text{tot}}$ and normalized excitations $\hat{e}_n = e_n/(c_{\text{tot}}\bar{S})$, since a_n is independent of c_{tot} and \bar{S} . The statistics of \hat{c}_i can be estimated in the typical case where odors are comprised of many ligands, see SI. In the particular case where the ligands are identically distributed the mean is $\langle \hat{c}_i \rangle = N_L^{-1}$ and the variance reads $\text{var}(\hat{c}_i) \approx (1 - p + \sigma^2 \mu^{-2})/(p N_L^2)$. Generally, \hat{c}_i varies more if the underlying c_i has higher coefficient of variation σ_i/μ_i or if the mean odor size $s = p N_L$ is smaller. The normalized excitation \hat{e}_n is defined such that its mean is 1 and the associated variance can be written as a product of the external contribution $V_{\text{ext}} = \sum_i \langle \hat{c}_i^2 \rangle$ due to odors and the internal contribution $V_{\text{int}} = \text{var}(S_{ni}) \langle S_{ni} \rangle^{-2}$ due to sensitivities, see SI. In the

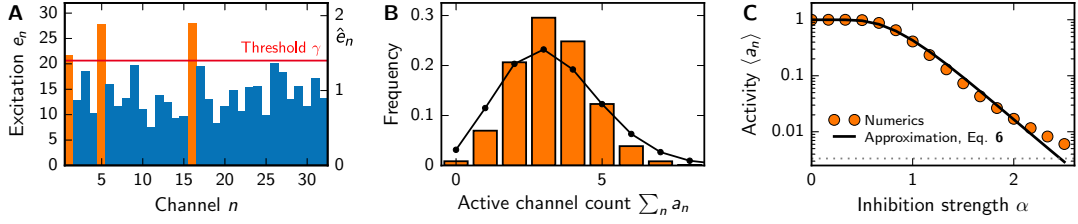


FIG. 2 Global inhibition with thresholding leads to sparse odor representations **a**. (A) Excitations e_n for an arbitrary odor. Active channels (orange) have an excitation above the threshold (red line, inhibition strength $\alpha = 1.4$). The right axis indicates the normalized excitation $\hat{e}_n = e_n N_R / \sum_m e_m$. (B) Histogram of the number of active channels compared to a binomial distribution (black line) with the same mean for $\alpha = 1.4$. (C) Mean channel activity $\langle a_n \rangle$ as a function of α . The approximation given by Eq. 6 (solid line) is compared to numerical simulations (symbols, standard error of the mean smaller than symbol size). The gray dotted line indicates a single expected active channel in humans, $\langle a_n \rangle = \frac{1}{300}$. (A–C) Additional model parameters are $N_R = 32$, $N_L = 256$, $p_i = 0.1$, $\mu_i = \sigma_i = 1$, and $\lambda = 1$.

simple case of identically distributed ligands, we have

$$\text{var}(\hat{e}_n) = V_{\text{ext}} V_{\text{int}} \quad V_{\text{ext}} \approx \frac{1}{s} \left(1 + \frac{\sigma^2}{\mu^2} \right) \quad V_{\text{int}} = e^{\lambda^2} - 1, \quad (5)$$

for $1 \ll s \ll N_L$, see SI. The normalized excitations thus vary more if odors contain fewer ligands, concentrations fluctuate stronger, or sensitivities are distributed more broadly. Finally, the mean channel activity $\langle a_n \rangle$ is given by the probability that the excitation e_n exceeds the threshold γ , see Eq. 2. This is equal to the probability that the normalized excitation \hat{e}_n exceeds the normalized threshold $\hat{\gamma} = \gamma / (\bar{S}_{\text{tot}})$. Replacing $\hat{\gamma}$ by its expectation value $\langle \hat{\gamma} \rangle = \alpha$ and using log-normally distributed e_n , we obtain

$$\langle a_n \rangle \approx \frac{1}{2} \text{erfc} \left(\frac{\zeta + \ln \alpha}{2\zeta^{\frac{1}{2}}} \right) \quad \text{with} \quad \zeta = \frac{1}{2} \ln(1 + V_{\text{ext}} V_{\text{int}}) \quad (6)$$

for log-normally distributed \hat{e}_n , see SI. Fig. 2C shows that this is a good approximation of the numerical results, which have been obtained from ensemble averages of Eq. 2.

The mean activity $\langle a_n \rangle$ can also be interpreted as the mean fraction of channels that are activated by an odor, such that small $\langle a_n \rangle$ corresponds to sparse odor representations. Fig. 2C shows that in our model this is the case for large inhibition strength α , where $\langle a_n \rangle \sim e^{-\nu}$ with $\nu \approx (\ln \alpha)^2 / (4\zeta)$, see SI. Since sparse representations are thought to be efficient for further processing in the brain (Laurent, 1999; Olshausen and Field, 2004) the inhibition strength α could be tuned, e.g., on evolutionary time scales, to achieve an activity $\langle a_n \rangle$ that is optimal for processing the odor representation downstream. If the optimal value of $\langle a_n \rangle$ is the same across animals, our theory predicts that inhibition is stronger in systems with more receptor types. However, this simple argument is not sufficient, since $\langle a_n \rangle$ also depends on the variations in the natural odor statistics and the

receptor sensitivities, which determine V_{ext} and V_{int} , respectively. In particular, the width λ of the sensitivity distribution could also be under evolutionary control. However, experimental data suggests that both flies and humans exhibit $\lambda \approx 1$ (Zwicker, Murugan, and Brenner, 2016). Additionally, we show in the SI that much smaller or larger values lead to extremely sparse representations, such that we will only consider $\lambda = 1$ in the following. In this case, the inhibition strength α controls the sparsity of the odor representation in our simple model of the olfactory system.

B. Sparse coding transmits useful information

One problem with sparse representations is that they cannot encode as many odors as dense representations. There is thus a maximal sparsity at which typical olfactory tasks can still be performed. In general, the performance of the olfactory system can be quantified by the transmitted information I , which is defined in Eq. 4. If we for simplicity neglect the small correlations between channels, I can be approximated as (Zwicker, Murugan, and Brenner, 2016)

$$I \approx - \sum_{n=1}^{N_R} [\langle a_n \rangle \log_2 \langle a_n \rangle + (1 - \langle a_n \rangle) \log_2 (1 - \langle a_n \rangle)] . \quad (7)$$

A maximum of N_R bits is transmitted when half the channels are active on average, $\langle a_n \rangle = \frac{1}{2}$. In our model, this is the case for weak inhibition, $\alpha < 1$, see Fig. 2C. In the opposite case of significant inhibition, $\alpha > 1$, few channels are typically active and the transmitted information is smaller. In the limit $\langle a_n \rangle \ll 1$, the information is approximately given by $I \sim \frac{1}{\ln 2} N_R \langle a_n \rangle \cdot (1 - \ln \langle a_n \rangle)$, which implies that even if only 10 % of the channels are active on average, the information I is still almost half of the maximal value of N_R bits. However, large information I does not automatically indicate a good receptor array, since only accessible information that can be used to solve a

given task matters (Tikhonov, Little, and Gregor, 2015; Tkačik and Bialek, 2016).

To test whether sparse representations are sufficient to solve typical olfactory tasks, we next study how well odors can be discriminated in our model. As a proxy for the discriminability, we calculate the Hamming distance d between the odor representations, which is given by the number of channels with different activity. In the simple case of uncorrelated odors, which do not share any ligands, the expected distance $\langle d \rangle$ is approximately given by total number of active channels in both representations. Consequently, uncorrelated odors can be distinguished even if their representations are very sparse. However, realistic tasks typically require distinguishing similar odors. We thus next study the discriminability of odors that vary in the relative concentrations of their ligands, their size, and their composition.

We start by determining the maximal dilution $\frac{c_b}{c_t}$ at which a target odor at concentration c_t can still be detected in a background of concentration c_b . We calculate the expected difference $\langle d \rangle$ between the associated representations from the probability that a given channel changes its activity when the target is added, see SI. Since this probability is the same for all channels, $\langle d \rangle$ is proportional to the number N_R of channels. For the simple case where both the target and the background are a single ligand, Fig. 3A shows that $\langle d \rangle$ decreases for smaller target concentrations and is qualitatively the same for all inhibition strengths α . For large dilutions $\frac{c_b}{c_t}$, $\langle d \rangle$ is inversely proportional to the dilution, $\langle d \rangle \propto N_R \frac{c_t}{c_b}$. Since the addition of the target can only be detected reliably if $\langle d \rangle > 2$, which corresponds to a situation where one channel becomes inactive and another one active, our model predicts that doubling the number N_R of channels also doubles the concentration sensitivity. Fig. 3A thus implies that mice ($N_R \approx 1000$) should be able to detect the addition of a target even if it is almost a hundred times more dilute than the background, which is close to the threshold that has been found experimentally (Mouret *et al.*, 2009). Conversely, flies ($N_R \approx 50$) should fail for very small dilution factors.

We next study odors comprised of many ligands, since typical odors are blends (Wright and Thomson, 2005). For simplicity, we consider the detection of a single target ligand in a background mixture of varying size s when the target ligand and the ligands in the background have equal concentration, such that the target dilution is s . Fig. 3B shows that the qualitative dependence of $\langle d \rangle$ on the dilution is similar to the single ligand case in panel A, but the maximal dilution for detecting the target is different. For instance, the model predicts that mice cannot identify the addition of the target ligand to a background consisting of more than ten ligands, while the maximal dilution was almost one hundred in the case of single background ligands. Consequently, the discrimination performance seems to drop significantly when larger odors are

considered. This qualitatively agrees with experiments where humans are not able to identify all ligands in mixtures of more than three ligands (Goyert *et al.*, 2007; Jinks and Laing, 2001) and they fail to detect the presence or absence of ligands in mixtures of more than 15 ligands (Jinks and Laing, 1999).

Even if humans cannot identify ligands in large odors, they might still be able to distinguish two such odors. To study this, we next compare the representations of two odors that each contain s ligands, sharing s_B of them, for the simple case where all ligands have the same concentration. Fig. 3C shows that the distance $\langle d \rangle$ between the two odors decreases with larger s_B , i.e., more similar odors are more difficult to discriminate. However, s_B only has a strong effect if more than about 80 % of the ligands are shared between odors. Conversely, the inhibition strength α and the odor size s significantly influence $\langle d \rangle$ for all values of s_B . This agrees with the results shown in Fig. 3B, where $\langle d \rangle$ exhibits a similar dependence on α and s . While it is expected that the performance decreases with large inhibition strength α since fewer channels are active, the strong dependence on the size s is surprising.

C. Larger odors have sparser representations

Why are odors with many ligands more difficult to discriminate in our model? Since correlations between channels seem to be negligible, the most likely explanation is that larger odors activate fewer channels. To test this hypothesis, we determine the activity $\langle a_n \rangle$ in the simple case where all ligands in an odor have the same concentration. Because of the normalization, the value of this concentration does not matter and $\langle a_n \rangle$ only depends on the inhibition strength α and the odor size s . In the limit of large odors ($s \gg 1$), the approximation given in Eq. 6 yields $\langle a_n \rangle \sim e^{-\beta s}$ with $\beta \sim (\ln \alpha)^2$, see SI. In this case, the activity $\langle a_n \rangle$ thus decreases exponentially with s and this decrease is stronger for larger α . Consequently, larger odors activate fewer channels and it is thus less likely that a small change in such odors alters the activation pattern \mathbf{a} .

Larger odors activate fewer channels because the respective excitations e_n have a smaller variability. For an odor with s ligands of equal concentration, e_n is proportional to the sum of s sensitivities S_{ni} , see Eq. 1. Consequently, e_n can be considered as a random variable whose mean $\langle e_n \rangle$ and variance $\text{var}(e_n)$ scale with s . The activity $\langle a_n \rangle$ is given by the fraction of excitations that exceed the threshold γ , which also scales with s . This fraction typically scales with the coefficient of variation $\text{var}(e_n)^{\frac{1}{2}} \langle e_n \rangle^{-1}$, which is proportional to $s^{-\frac{1}{2}}$ and is thus smaller for larger odors. Larger odors thus activate fewer channels because there are fewer excitations that are much larger than the mean, see Fig. 4A.

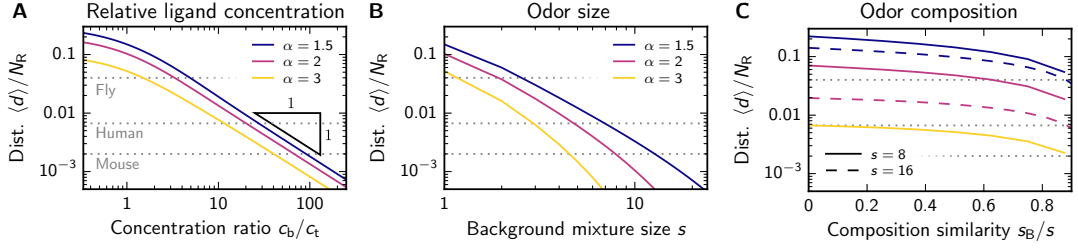


FIG. 3 Sparse coding is sufficient to distinguish odors with different relative ligand concentrations, size, and composition. (A) Mean distance $\langle d \rangle$ between the representations of a background ligand at concentration c_b and an odor with an additional target ligand at concentration c_t as a function of the dilution c_b/c_t for various inhibition strengths α . (B) Distance $\langle d \rangle$ resulting from adding a ligand to an odor comprised of s ligands as a function of s for various α . (C) Distance $\langle d \rangle$ between the representations of two odors with s ligands, sharing s_B of them, as a function of the similarity s_B/s for small ($s = 8$, solid lines) and large odors ($s = 16$, dashed lines). The colors indicate the same α as in the other panels. (A–C) The gray dotted lines indicate the threshold $\langle d \rangle = 2$ for $N_R = 50, 300, 1000$ (corresponding to flies, humans, and mice; top to bottom). The width of the sensitivity distribution is $\lambda = 1$.

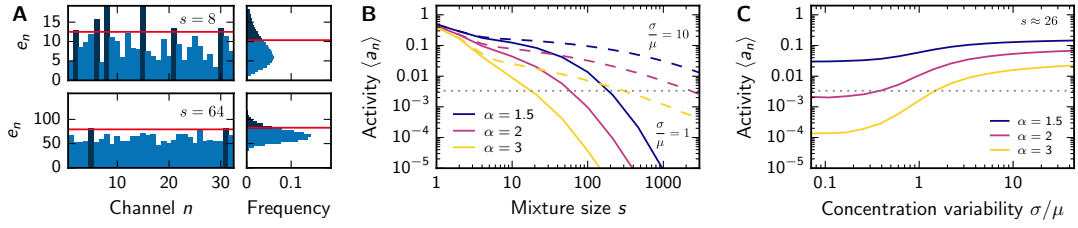


FIG. 4 Larger odors activate fewer channels. (A) Comparison of the excitations e_n of small (odor size $s = 8$, upper panels) and large odors ($s = 64$, lower panels) at $\alpha = 1.3$. e_n for a single odor (left panels) and histograms for all odors (right panels) are shown. Larger odors exhibit fewer active channels (dark blue), for which the excitations are above threshold (red line). (B) Numerically determined $\langle a_n \rangle$ as a function of s for various inhibition strengths α at small ($\sigma/\mu = 1$, solid lines) and large concentration variability ($\sigma/\mu = 10$, dashed lines) at $N_L = 10^4$. (C) Numerically determined $\langle a_n \rangle$ as a function of σ/μ for various α . (A–C) Additional model parameters are $N_R = 32$, $N_L = 256$, $p_i = 0.1$, $\mu_i = \sigma_i = 1$, and $\lambda = 1$. The gray dotted line in B and C indicates a single expected active channel in humans, $\langle a_n \rangle = \frac{1}{300}$.

This is a direct consequence of the assumption that the excitation threshold γ scales with the mean excitation and this result does not depend on other details of the model. Conversely, the dependence of $\langle a_n \rangle$ on the inhibition strength α is model specific, since it follows from the shape of the tail of the excitation distribution. In particular, the influence of the odor size on $\langle a_n \rangle$ is insignificant for weak inhibition, $\alpha \approx 1$, because approximately half the channels are activated irrespective of the variance $\text{var}(e_n)$.

This qualitative explanation illustrates that depending on the variability of the excitations different odors can have representations with very different sparsities. Indeed, we find that the sparsity changes over several orders of magnitude as a function of the odor size s in our model, see Fig. 4B. Moreover, the concentration variability $\frac{\sigma}{\mu}$ of the individual ligands also has a strong effect on the sparsity, see Fig. 4C. This is because larger $\frac{\sigma}{\mu}$ implies larger variations in the excitations, such that more channels exceed the threshold and become active. In fact, this dependence of $\langle a_n \rangle$ on s and $\frac{\sigma}{\mu}$ is also qualitatively captured by the analytical approximation given in Eq. 6, which explicitly depends on the odor variability V_{ext} defined in Eq. 5. Taken together, our model shows that the

sparsity of the odor representations strongly depend on the odor statistics $P_{\text{env}}(\mathbf{c})$.

D. Effective arrays have similar receptor sensitivities

So far, we considered homogeneous receptor arrays, where all receptor types have the same average sensitivity. However, realistic receptors vary in their biochemical details and it might thus be difficult to have such homogeneous arrays. We thus next consider the effect of sensitivity variations between different receptors. This is important, since a channel with overly sensitive receptors will contribute significantly to the common threshold γ , suppress the activity of other channels, and could thus limit the coding capacity of the system, see Fig. 5A. To study this, we consider sensitivity matrices $S_{ni} = \xi_n S_{ni}^{\text{id}}$, where ξ_n denotes the mean sensitivity of receptor type n and S_{ni}^{id} is the sensitivity matrix that we discussed so far, i.e., it is a random matrix where all entries are independently drawn from a log-normal distribution described by the mean \bar{S} and width λ . Here, ξ_n captures differences between receptor types, e.g., because of biochemical differences or due to variations in copy number,

see SI. For this model, the mean excitation threshold is $\langle\gamma\rangle = \alpha\bar{S}\langle c_{\text{tot}}\rangle\xi_{\text{tot}}N_R^{-1}$ where $\xi_{\text{tot}} = \sum_n \xi_n$. The expected channel activity is approximately given by

$$\langle a_n \rangle \approx 1 - F\left(\frac{\alpha\xi_{\text{tot}}}{N_R\xi_n}\right), \quad (8)$$

where $F(\hat{e}_n)$ is the cumulative distribution function of the normalized excitations \hat{e}_n for $\xi_n = 1$, whose mean is $\langle\hat{e}_n\rangle = 1$ and whose variance is given by Eq. 5. Note that $\langle a_n \rangle$ does not change if all ξ_n are multiplied by the same factor. In particular, the expression above reduces to $\langle a_n \rangle \approx 1 - F(\alpha)$ and thus Eq. 6 if all ξ_n are equal.

We first discuss the influence of the receptor sensitivities ξ_n by only varying one type, i.e., we change ξ_1 while setting $\xi_n = 1$ for $n \geq 2$. Fig. 5B shows that for fixed channel activity $\langle a_n \rangle$ the transmitted information I is maximal for a homogeneous receptor array ($\xi_1 = 1$). I is reduced for smaller ξ_1 and for $\xi_1 = 0$ it reaches the value I_0 of an array where the first receptor was removed. Conversely, I can drop well below I_0 when ξ_1 is increased above 1. In this case, the large excitation of the affected channel not only leads to its likely activation, but it also raises the threshold γ and thereby inhibits other channels, see Fig. 5A. In the extreme case of very large ξ_1 , this channel will always be active while all other channels are silenced, which implies $I = 0$. There is thus a critical value of ξ_1 beyond which removing the receptor from the array is advantageous for the overall performance. Fig. 5B shows that increasing the sensitivity of a receptor by only 40% can make it useless in the context of the whole array if representations are sparse.

So far, we only varied the sensitivity of a single receptor. To test how variations in the sensitivities of all receptors affect the information I , we next consider log-normally distributed ξ_n . Here, vanishing variance of ξ_n corresponds to a homogeneous receptor array. Fig. 5C shows that small variations in ξ_n can strongly reduce the transmitted information I . Since I limits the discriminative capability of the receptor array, this suggests that receptor arrays with heterogeneous sensitivities perform worse.

The simple model that we discuss here shows that the excitation statistics of the different channels determine the properties of the resulting odor representation. In particular, receptors that have lower excitations on average might be suppressed often and thus contribute less to the odor information. Since the excitation statistics are influenced both by the sensitivities S_{ni} and the odor statistics $P_{\text{env}}(\mathbf{c})$, this suggests that the sensitivities should be adjusted to the odor statistics. In an optimal receptor array, the sensitivities are chosen such that all channels have the same probability to become active.

IV. DISCUSSION

We studied a simple model of odor representations, which is based on normalization and a non-linear gain function. This model separates the odor composition, encoded in the activity \mathbf{a} of the projection neurons, from the odor intensity, which could be encoded by the total excitation e_{tot} or the threshold level γ (Mainland *et al.*, 2014). For significant inhibition the representation \mathbf{a} is sparse and the set of active projection neurons provides a natural odor 'tag' that could be used for identification and memorization in the downstream processing (Stevens, 2015).

Sparse representations reduce the coding capacity and transmit less information than dense ones. However, even if the mean activity is $\langle a_n \rangle = 0.01$ and thus 50 times smaller than in maximally informative arrays with $\langle a_n \rangle = 0.5$, the transmitted information I is only reduced by a factor of 12, see Eq. 7. For humans with $N_R = 300$, this yields $I \approx 25$ bits, allowing to encode $2^I \approx 10^7$ different odor compositions. Note that the total information I_{tot} also includes information I_{int} about the odor intensity, $I_{\text{tot}} = I + I_{\text{int}}$. Here, $I_{\text{int}} \approx 10$ bits would be sufficient to encode the total concentration over a range of 10 orders of magnitude with a resolution of 5%, typical for humans (Cain, 1977). In this case, our model compresses the 300 bits of a maximally informative representation on the level of glomeruli (Zwicker, Murugan, and Brenner, 2016) to only $I_{\text{tot}} \approx 35$ bits on the level of projection neurons.

The model discussed here is similar to our previous model, where we discussed representations on the level of the glomeruli (Zwicker, Murugan, and Brenner, 2016). Both models use a maximum entropy principle to determine properties of optimal receptor arrays. To achieve this, the receptor sensitivities must be tailored to the odor statistics in both models. The main difference of the models is the global inhibition discussed here, which separates the odor composition from its intensity and thus removes the correlation between the glomeruli excitation and the odor intensity (Haddad *et al.*, 2010). Consequently, odors can then be discriminated at all concentrations, while this was only possible in a narrow concentration range in the glomeruli model (Zwicker, Murugan, and Brenner, 2016). The additional normalization is thus useful to separate odors, even if the projection neurons encode less information than the respective glomeruli. To estimate this information, we consider binary outputs in both models, which corresponds to very noisy channels. However, the glomeruli model discusses arrays of noisy receptor, while we here consider perfect receptors whose signal is first normalized and then subjected to noise. This additional processing reduces correlations and leads to sparse representations, which might simplify downstream computations. Consequently, this model is suitable for describing natural olfaction, where the capac-

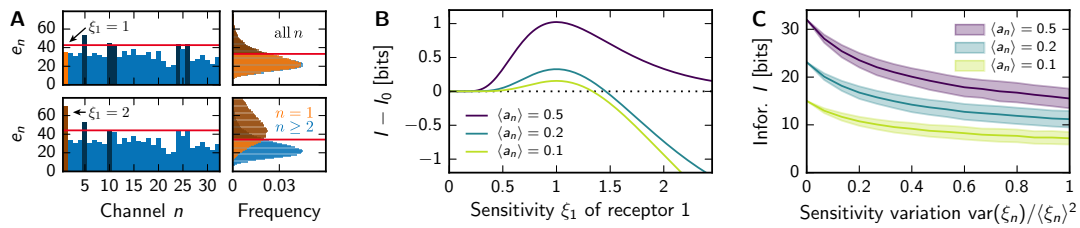


FIG. 5 Receptors with heterogeneous sensitivities make poor arrays. (A) Comparison of the excitations e_n for homogeneous ($\xi_1 = 1$, upper panels) and heterogeneous receptors ($\xi_1 = 2$, lower panels). e_n for the same arbitrary odor (left panels) and histograms for all odors (right panels) are shown for the first receptor ($n = 1$, orange) and all other receptors ($n \geq 2$, blue). Dark bars indicate excitations that are above the threshold (red line, inhibition strength $\alpha = 1.3$). (B) Information I given by Eq. 7 as a function of the sensitivity ξ_1 of the first receptor. The channel activity $\langle a_n \rangle$ calculated from Eq. 8 is set to the given value by adjusting α . I is shown relative to the information I_0 of a system without the first receptor (dotted line). (C) Information I (line, mean; shaded area indicates standard deviation) of log-normally distributed ξ_n as a function of the variation $\text{var}(\xi_n)\langle \xi_n \rangle^{-2}$ for various $\langle a_n \rangle$. (A–C) Remaining parameters are $N_R = 32$, $N_L = 256$, $p_i = 0.1$, $\mu_i = \sigma_i = 1$, and $\lambda = 1$.

ity for the downstream computations is limited, while the glomeruli model is relevant for artificial olfaction (Stitzel, Aernecke, and Walt, 2011), since computers have enough power to handle high-dimensional signals.

Sparse responses of projection neurons have been observed in experiments (Davison and Katz, 2007; Rinberg, Koulakov, and Gelperin, 2006). For instance, in mice 15 % of the projection neurons respond to a given single ligand (Roland *et al.*, 2016), suggesting significant inhibition. However, in locust about two third of the projection neurons respond to any given odor (Perez-Orive *et al.*, 2002), which implies weak inhibition. It is thus conceivable that some animals exhibit sparse representations while others have maximally informative ones, although additional experiments are needed to characterize the representations better. A direct experiment could test whether the odor percept changes when the weakly responding glomeruli are disabled artificially. Additionally, it will be important to study the representations of mono-molecular odors and mixtures at various concentration to better resemble the natural odor statistics. For instance, our simple theory predicts that fewer than 15 % of the projection neurons in mice respond when complex mixtures are presented. Indeed, experiments find that only 3 to 10 % of the projection neurons in mice fire for complex urine odors (Lin *et al.*, 2005). Conversely, the statistics of the activity of projection neurons in flies seem to be independent of the stimulus (Stevens, 2016). Our theory can also be tested by measuring how well odors can be discriminated. For instance, odors are much more difficult to distinguish if they contain more ligands in our model, which has also been observed experimentally (Weiss *et al.*, 2012). Conversely, other experiments indicate that the odor size only weakly influences the odor discriminability (Bushdid *et al.*, 2014). Taken together, there is some experimental evidence that the odor representations and thus the discriminability change with odor size, although there is also evidence to the contrary, which could hint at mechanisms beyond global inhibition

that influence the odor representations.

The coding sparsity given by the mean channel activity $\langle a_n \rangle$ can be adjusted by changing the inhibition strength α or the width λ of the receptor sensitivity distribution in our model. Additionally, $\langle a_n \rangle$ is a function of the natural odor statistics, i.e., the typical number of ligands in odors and their concentration distribution. Consequently, α or λ must be adjusted to keep $\langle a_n \rangle$ constant if the odor statistics change, e.g., because of seasonal changes or migration to a different environment. This adjustment could happen on multiple timescales, reaching from evolutionary adaptations of the receptors to near-instantaneous adjustments of the involved neurons, and it is likely that the global inhibition is regulated on all levels (Wilson, 2013). In this paper, we investigated the simple case of constant α and λ , which corresponds to slow regulation, but it is conceivable that α could be regulated on short time scales. For instance, the threshold could be lowered for larger odors to improve their discriminability. Our model suggests that such additional mechanisms are necessary to efficiently discriminate odors of all sizes.

Our model also reveals that it is important to control the properties of the individual communication channels to have useful receptor arrays. For instance, increasing the sensitivity of a given receptor by 40 % can be worse than removing it completely, see Fig. 5A. Generally, a receptor array is only effective if the different channels have similar excitations on average. This suggests that the sensitivities are tightly controlled and maybe even adjusted to the odor statistics of the environment. On evolutionary time scales, the sensitivities could be regulated by point mutations of the receptors that change how ligands bind (Yu *et al.*, 2015). On shorter time scales, the sensitivities could be regulated by changing the receptor copy numbers, see SI. Since this is observed experimentally (Yu and Wu, 2016), we predict that the receptor copy numbers are adjusted such that the excitations of all glomeruli are similar when averaged over

natural odors. Alternatively, variations in the receptor sensitivities could be balanced by more complex inhibition mechanism. For instance, experiments show that different projection neurons have different susceptibilities to inhibition (Hong and Wilson, 2015). Here, the experimentally observed turnover of mitral cells and interneurons (Lazarini and Lledo, 2011) could adjust the inhibition mechanism locally, which could optimize the olfactory system for a given environment (Mouret *et al.*, 2009). Such adaptation of the inhibition mechanism to the current stimulus statistics and more complex models where the behavioral state of an animal could influence the olfactory bulb by top-down modulation (Wilson, 2013) will be interesting to explore in the future.

Our simplified model neglects many details of the olfactory system (Silva Teixeira, Cerqueira, and Silva Ferreira, 2016). For instance, we do not consider the dynamics of inhalation and the odor absorption in the mucus (Pelosi, 2001; Schoenfeld and Cleland, 2005). Instead, we here directly parameterize the ligand distribution at the olfactory receptors, where we for simplicity neglect correlations between ligands. It would be interesting to extend the model for more complex stimuli and study how the system decorrelates the input, identifies a target odor in a background, and separates multiple odors from each other. This likely involves many steps (Cleland *et al.*, 2011) and cannot be done perfectly with a single normalization step and non-linear gain function. For instance, it might be important to apply gain functions at the level of receptors and the glomeruli to model finite sensitivity and saturation effects. Additionally, it has been shown that there is additional cross-talk on the level of receptors (Ukhanov *et al.*, 2010) and glomeruli (Aungst *et al.*, 2003; Silbering and Galizia, 2007), which could support decorrelation. Generally, such cross-talk and the inhibition that we discussed here will be non-linear (Wilson, 2011). This could for instance be modeled by a divisive normalization model that has been proposed for olfaction (Olsen, Bhandawat, and Wilson, 2010). It is also likely that the inhibition of the projection neurons is not driven by a single global variable. If glomeruli positioning carried some meaning (Murthy, 2011), local inhibition could help separating similar odors by enhancing the contrast (Leon and Johnson, 2003). The discrimination of similar odors could also be improved if projection neurons had a larger output range, increasing the information capacity per channel. Finally, we completely neglected the temporal dynamics of the olfactory system, which play an important role for the adaptation between sniffs (Zufall and Leinders-Zufall, 2000) and might also influence odor perception within a single sniff (Blauvelt *et al.*, 2013; Sirotin, Shusterman, and Rinberg, 2015; Uchida, Poo, and Haddad, 2014).

ACKNOWLEDGMENTS

I thank Michael P. Brenner, Venkatesh N. Murthy, Mikhail Tikhonov and Christoph A. Weber for helpful discussions and a critical reading of the manuscript. This research was funded by the Simons Foundation and the German Science Foundation through ZW 222/1-1.

REFERENCES

- Asahina, K., Louis, M., Piccinotti, S., and Vosshall, L. B., *J Biol* **8**, 9 (2009).
- Aungst, J. L., Heyward, P. M., Puche, A. C., Karnup, S. V., Hayar, A., Szabo, G., and Shipley, M. T., *Nature* **426**, 623 (2003).
- Banerjee, A., Marbach, F., Anselmi, F., Koh, M. S., Davis, M. B., Garcia da Silva, P., Delevich, K., Oyibo, H. K., Gupta, P., Li, B., and Albeanu, D. F., *Neuron* **87**, 193 (2015).
- Barlow, H., *Network* **12**, 241 (2001).
- Barlow, H. B., in *Sensory Communication*, edited by W. Rosenblith (MIT press, 1961) pp. 217–234.
- Berck, M. E., Khandelwal, A., Claus, L., Hernandez-Nunez, L., Si, G., Tabone, C. J., Li, F., Truman, J. W., Fetter, R. D., Louis, M., Samuel, A. D., and Cardona, A., *Elife* **5** (2016), 10.7554/eLife.14859.
- Bhandawat, V., Olsen, S. R., Gouwens, N. W., Schlieff, M. L., and Wilson, R. I., *Nat. Neurosci.* **10**, 1474 (2007).
- Blauvelt, D. G., Sato, T. F., Wienisch, M., and Murthy, V. N., *Frontiers in neural circuits* **7** (2013).
- Bushdid, C., Magnasco, M., Vosshall, L., and Keller, A., *Science* **343**, 1370 (2014).
- Cain, W. S., *Science* **195**, 796 (1977).
- Carandini, M. and Heeger, D. J., *Nat Rev Neurosci* **13**, 51 (2012).
- Cleland, T. A., *Trends. Neurosci.* **33**, 130 (2010).
- Cleland, T. A., Chen, S.-Y. T., Hozer, K. W., Ukatu, H. N., Wong, K. J., and Zheng, F., *Front Neuroeng* **4**, 21 (2011).
- Cleland, T. A., Johnson, B. A., Leon, M., and Linster, C., *Proc. Natl. Acad. Sci. USA* **104**, 1953 (2007).
- Cleland, T. A. and Sethupathy, P., *BMC Neurosci* **7**, 7 (2006).
- Davison, I. G. and Katz, L. C., *J Neurosci* **27**, 2091 (2007).
- Demb, J. B. and Singer, J. H., *Annual Review of Vision Science* **1**, 263 (2015).
- Getz, W. M. and Lutz, A., *Chem. Senses* **24**, 351 (1999).
- Goyert, H. F., Frank, M. E., Gent, J. F., and Hettinger, T. P., *Brain Res Bull* **72**, 1 (2007).
- Gupta, P., Albeanu, D. F., and Bhalla, U. S., *Nat. Neurosci.* **18**, 272 (2015).
- Haddad, R., Weiss, T., Khan, R., Nadler, B., Mandairon, N., Bensafi, M., Schneidman, E., and Sobel, N., *J Neurosci* **30**, 9017 (2010).
- Hong, E. J. and Wilson, R. I., *Neuron* **85**, 573 (2015).
- Hopfield, J., *Proc. Natl. Acad. Sci. USA* **96**, 12506 (1999).
- Isaacson, J. S. and Scanziani, M., *Neuron* **72**, 231 (2011).
- Jefferis, G. S., Marin, E. C., Stocker, R. F., and Luo, L., *Nature* **414**, 204 (2001).
- Jinks, A. and Laing, D. G., *Perception* **28**, 395 (1999).
- Jinks, A. and Laing, D. G., *Physiol Behav* **72**, 51 (2001).
- Kaupp, U. B., *Nat Rev Neurosci* **11**, 188 (2010).
- Koulakov, A., Gelperin, A., and Rinberg, D., *J. Neurophysiol.* **98**, 3134 (2007).

- Laurent, G., *Science* **286**, 723 (1999).
- Lazarini, F. and Lledo, P.-M., *Trends Neurosci* **34**, 20 (2011).
- Leon, M. and Johnson, B. A., *Brain Res. Rev.* **42**, 23 (2003).
- Li, Z., *Biol Cybern* **62**, 349 (1990).
- Li, Z., in *Models of neural networks* (Springer, 1994) Chap. 6, pp. 221–251.
- Lin, D. Y., Zhang, S.-Z., Block, E., and Katz, L. C., *Nature* **434**, 470 (2005).
- Linstner, C. and Hasselmo, M., *Behavioural brain research* **84**, 117 (1997).
- Lowe, G. and Gold, G. H., *Proc. Natl. Acad. Sci. USA* **92**, 7864 (1995).
- Mainland, J. D., Lundström, J. N., Reiser, J., and Lowe, G., *Trends Neurosci.* **37**, 443 (2014).
- Malnic, B., Hirono, J., Sato, T., and Buck, L. B., *Cell* **96**, 713 (1999).
- Mouret, A., Lepousez, G., Gras, J., Gabellec, M.-M., and Lledo, P.-M., *J Neurosci* **29**, 12302 (2009).
- Murthy, V. N., *Annu. Rev. Neurosci.* **34**, 233 (2011).
- Niimura, Y., *Curr Genomics* **13**, 103 (2012).
- Nikolova, N. and Jaworska, J., *QSAR & Combinatorial Science* **22**, 1006 (2003).
- Olsen, S. R., Bhandawat, V., and Wilson, R. I., *Neuron* **66**, 287 (2010).
- Olshausen, B. A. and Field, D. J., *Curr Opin Neurobiol* **14**, 481 (2004).
- Pelosi, P., *Cellular and Molecular Life Sciences CMLS* **58**, 503 (2001).
- Perez-Orive, J., Mazor, O., Turner, G. C., Cassenaer, S., Wilson, R. I., and Laurent, G., *Science* **297**, 359 (2002).
- Rinberg, D., Koulakov, A., and Gelperin, A., *J Neurosci* **26**, 8857 (2006).
- Roland, B., Jordan, R., Sosulski, D. L., Diodato, A., Fukunaga, I., Wickersham, I., Franks, K. M., Schaefer, A. T., and Fleischmann, A., *Elife* **5** (2016), 10.7554/eLife.16335.
- Ruderman, and Bialek, *Phys. Rev. Lett.* **73**, 814 (1994).
- Sachse, S. and Galizia, C. G., *J Neurophysiol* **87**, 1106 (2002).
- Schoenfeld, T. A. and Cleland, T. A., *Trends in neurosciences* **28**, 620 (2005).
- Silbering, A. F. and Galizia, C. G., *J. Neurosci.* **27**, 11966 (2007).
- Silva Teixeira, C. S., Cerqueira, N. M. F. S. A., and Silva Ferreira, A. C., *Chem Senses* **41**, 105 (2016).
- Sirotin, Y. B., Shusterman, R., and Rinberg, D., *eNeuro* **2** (2015), 10.1523/ENEURO.0083-15.2015.
- Soucy, E. R., Albeanu, D. F., Fantana, A. L., Murthy, V. N., and Meister, M., *Nat. Neurosci.* **12**, 210 (2009).
- Stevens, C. F., *Proc. Natl. Acad. Sci. USA* **112**, 9460 (2015).
- Stevens, C. F., *Proc. Natl. Acad. Sci. USA* (2016), 10.1073/pnas.1606339113.
- Stitzel, S. E., Aernicke, M. J., and Walt, D. R., *Annu. Rev. Biomed. Eng.* **13**, 1 (2011).
- Su, C.-Y., Menuz, K., and Carlson, J. R., *Cell* **139**, 45 (2009).
- Tabor, R., Yaksi, E., Weislogel, J.-M., and Friedrich, R. W., *J. Neurosci.* **24**, 6611 (2004).
- Tan, J., Savigner, A., Ma, M., and Luo, M., *Neuron* **65**, 912 (2010).
- Tikhonov, M., Little, S. C., and Gregor, T., *R Soc Open Sci* **2**, 150486 (2015).
- Tkačik, G. and Bialek, W., *Annual Review of Condensed Matter Physics* **7**, 89 (2016), <http://dx.doi.org/10.1146/annurev-conmatphys-031214-014803>.
- Uchida, N. and Mainen, Z. F., *Front Syst Neurosci* **1**, 3 (2007).
- Uchida, N., Poo, C., and Haddad, R., *Annu Rev Neurosci* **37**, 363 (2014).
- Ukhanov, K., Corey, E. A., Brunert, D., Klasen, K., and Ache, B. W., *J Neurophysiol* **103**, 1114 (2010).
- Verbeurg, C., Wilkin, F., Tarabichi, M., Gregoire, F., Dumont, J. E., and Chatelain, P., *PLOS ONE* **9**, e96333 (2014).
- Weiss, T., Snitz, K., Yablonka, A., Khan, R. M., Gafsou, D., Schneidman, E., and Sobel, N., *Proc. Natl. Acad. Sci. USA* **109**, 19959 (2012).
- Wilson, R. I., *Curr. Opin. Neurobiol.* **21**, 254 (2011).
- Wilson, R. I., *Annu. Rev. Neurosci.* **36**, 217 (2013).
- Wright, G. A. and Thomson, M. G., in *Integrative Plant Biochemistry*, Recent Advances in Phytochemistry, Vol. 39, edited by J. Romeo (Elsevier, 2005) Chap. 8, pp. 191–226.
- Yokoi, M., Mori, K., and Nakanishi, S., *Proc. Natl. Acad. Sci. USA* **92**, 3371 (1995).
- Yu, C. R. and Wu, Y., *Exp Neurol* (2016), 10.1016/j.expneurol.2016.06.001.
- Yu, Y., Claire, A., Ni, M. J., Adipietro, K. A., Golebiowski, J., Matsunami, H., and Ma, M., *Proc. Natl. Acad. Sci. USA* **112**, 14966 (2015).
- Zhang, D., Li, Y., and Wu, S., *Comput Math Methods Med* **2013**, 507143 (2013).
- Zufall, F. and Leinders-Zufall, T., *Chem. Senses* **25**, 473 (2000).
- Zwicker, D., Murugan, A., and Brenner, M. P., *Proc. Natl. Acad. Sci. USA* **113**, 5570 (2016).

Supporting Information: Normalized neural representations of natural odors

David Zwicker^{1,2,*}

¹*School of Engineering and Applied Sciences, Harvard University, Cambridge, MA 02138, USA*

²*Kavli Institute for Bionano Science and Technology, Harvard University, Cambridge, MA 02138, USA*

(Dated: August 4, 2016)

CONTENTS

S1. Statistics of normalized concentrations and excitations	1
S2. Numerical simulations	2
S3. Approximate channel activity	2
S4. Odor discriminability	2
S5. Receptor binding model	3

S1. STATISTICS OF NORMALIZED CONCENTRATIONS AND EXCITATIONS

Let p_i be the probability that ligand i is present in an odor. If it is present, its concentration c_i is drawn from a log-normal distribution with mean μ_i and standard deviation σ_i , while $c_i = 0$ if the ligand is not present. Hence,

$$\langle c_i \rangle = p_i \mu_i \quad (\text{S1a})$$

$$\text{var}(c_i) = (p_i - p_i^2) \mu_i^2 + p_i \sigma_i^2, \quad (\text{S1b})$$

while the covariances $\text{cov}(c_i, c_j) = \langle c_i c_j \rangle - \langle c_i \rangle \langle c_j \rangle$ vanish for $i \neq j$ since the ligands are independent. The statistics of the total concentration $c_{\text{tot}} = \sum_i c_i$ read

$$\langle c_{\text{tot}} \rangle = \sum_{i=1}^{N_L} \langle c_i \rangle \quad \text{and} \quad \text{var}(c_{\text{tot}}) = \sum_{i=1}^{N_L} \text{var}(c_i). \quad (\text{S2})$$

The excitations e_n are given by $e_n = \sum_i S_{ni} c_i$, where the sensitivities S_{ni} are log-normally distributed with mean $\langle S_{ni} \rangle = \bar{S}$ and variance $\text{var}(S_{ni}) = \bar{S}^2 (e^{\lambda^2} - 1)$. Hence,

$$\langle e_n \rangle = \bar{S} \langle c_{\text{tot}} \rangle \quad (\text{S3a})$$

$$\text{var}(e_n) = \bar{S}^2 \text{var}(c_{\text{tot}}) + \text{var}(S_{ni}) \sum_{i=1}^{N_L} \langle c_i^2 \rangle, \quad (\text{S3b})$$

where $\langle c_i^2 \rangle = p_i (\mu_i^2 + \sigma_i^2)$ and $\text{cov}(e_n, e_m) = 0$ for $n \neq m$.

We next determine the statistics of the normalized concentrations $\hat{c}_i = c_i / c_{\text{tot}}$. For simplicity, we consider large odors, $\sum_i p_i \gg 1$, where c_{tot} can be considered as an independent random variable. Since c_{tot} is the sum of

(a variable) number of log-normally distributed random variables, its distribution can be approximated by another log-normal distribution (Wu, Mehta, and Zhang, 2005), which we parameterize by its mean μ_{tot} and variance σ_{tot}^2 . We consider the simple approximation where these parameters are directly given by Eq. S2 (Fenton, 1960). This choice approximates the tail of the distribution well, but leads to errors in the vicinity of the mean (Wu, Mehta, and Zhang, 2005).

Since both c_{tot} and c_i are log-normally distributed when ligand i is present in an odor ($c_i > 0$), \hat{c}_i is also log-normally distributed in this case and

$$\langle \hat{c}_i \rangle_{c_i > 0} = \frac{\mu_i}{\mu_{\text{tot}}} \chi \quad (\text{S4a})$$

$$\text{var}(\hat{c}_i)_{c_i > 0} = \frac{\mu_i^2 \chi^2}{\mu_{\text{tot}}^2} \left(\frac{\sigma_i^2}{\mu_i^2} \chi + \chi - 1 \right), \quad (\text{S4b})$$

where $\chi = 1 + \sigma_{\text{tot}}^2 \mu_{\text{tot}}^{-2}$. Since $\hat{c}_i = 0$ with probability $1 - p_i$, the statistics of \hat{c}_i read

$$\langle \hat{c}_i \rangle = \frac{p_i \mu_i}{\mu_{\text{tot}}} \chi \quad (\text{S5a})$$

$$\text{var}(\hat{c}_i) = \frac{p_i \mu_i^2 \chi^2}{\mu_{\text{tot}}^2} \left(\frac{\sigma_i^2}{\mu_i^2} \chi + \chi - p_i \right). \quad (\text{S5b})$$

Note that the covariance $\text{cov}(\hat{c}_i, \hat{c}_j)$ does not vanish since the \hat{c}_i are not independent. In particular, $\text{var}(\sum_i \hat{c}_i) = 0$, since $\sum_i \hat{c}_i = 1$ by definition. This condition is only consistent with Eq. S5a if $\chi \approx 1$, which implies that c_{tot} must not vary much, $\frac{\sigma_{\text{tot}}}{\mu_{\text{tot}}} \ll 1$. Using $\chi = 1$, the statistics of the normalized excitations $\hat{e}_n = \bar{S}^{-1} \sum_i S_{ni} \hat{c}_i$ read

$$\langle \hat{e}_n \rangle = 1 \quad (\text{S6a})$$

$$\text{var}(\hat{e}_n) = \frac{\text{var}(S_{ni})}{\bar{S}^2} \left\langle \sum_i \hat{c}_i^2 \right\rangle, \quad (\text{S6b})$$

where $\langle \sum_i \hat{c}_i^2 \rangle \approx \sum_i \langle \hat{c}_i^2 \rangle$ with $\langle \hat{c}_i^2 \rangle = \langle \hat{c}_i \rangle^2 + \text{var}(\hat{c}_i)$ and the statistics given in Eq. S5.

In the simple case where all ligands are drawn from the same distribution ($p_i = p$, $\mu_i = \mu$, $\sigma_i = \sigma$), we obtain

$$\langle \hat{c}_i \rangle \approx \frac{1}{N_L}, \quad \text{var}(\hat{c}_i) \approx \frac{1 - p + \frac{\sigma^2}{\mu^2}}{s N_L}, \quad (\text{S7})$$

and $\langle \hat{c}_i^2 \rangle \approx \frac{1}{s N_L} (\frac{\sigma^2}{\mu^2} + 1)$, such that

$$\text{var}(\hat{e}_n) \approx \frac{1}{s} \left(1 + \frac{\sigma^2}{\mu^2} \right) \frac{\text{var}(S_{ni})}{\bar{S}^2}, \quad (\text{S8})$$

which is equivalent to Eq. 5 in the main text.

* dzwicker@seas.harvard.edu; <http://www.david-zwicker.de>

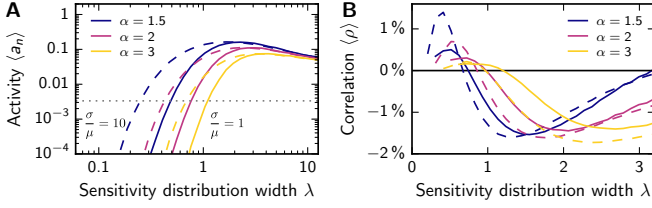


FIG. S1 Influence of the width λ of the sensitivity distribution on the statistics of the odor representations. (A) Expected channel activity $\langle a_n \rangle$ as a function of λ for several inhibition strengths α . Intermediated values, $\lambda \approx 1$, lead to larger activities. (B) Mean Pearson correlation coefficient $\langle \rho \rangle$ calculated from an ensemble average of Eq. S9 as a function of λ for several α . For small λ , $\langle a_n \rangle$ was too small to estimate ρ reliably. (A–B) Results are shown for small ($\frac{\sigma}{\mu} = 1$, solid lines) and large ($\frac{\sigma}{\mu} = 10$, dashed lines) concentration variability. Remaining parameters are $N_R = 32$, $N_L = 256$, and $p_i = 0.1$.

S2. NUMERICAL SIMULATIONS

We numerically calculated ensemble averages over odors \mathbf{c} and sensitivity matrices S_{ni} . Here, we first choose S_{ni} by drawing all entries independently from a log-normal distribution with mean $\bar{S} = 1$ and variance $\text{var}(S_{ni}) = e^{\lambda^2} - 1$. We then draw an odor \mathbf{c} using the following procedure: First, we determine which of the N_L ligands are present according to their probabilities p_i . Second, we draw the concentrations c_i for each ligand i that is present from a log-normal distribution with mean μ_i and standard deviation σ_i . We then use Eqs. 1–3 given in the main text to map the odor \mathbf{c} to a binary activity vector \mathbf{a} , from which we can for instance calculate the number of active channels. We obtain ensemble averages of such quantities by repeating these steps 10^5 times. This allows us to calculate the mean activities $\langle a_n \rangle$, the covariances $\text{cov}(a_n, a_m)$, and the Pearson correlation coefficient ρ , which is defined as

$$\rho = \frac{1}{N_R^2 - N_R} \sum_{n \neq m} \frac{\text{cov}(a_n, a_m)}{[\text{var}(a_n) \text{var}(a_m)]^{\frac{1}{2}}}. \quad (\text{S9})$$

Fig. S1 shows these quantities as a function of the width λ of the sensitivity distribution. We also estimate $P(\mathbf{a})$ from an ensemble average to calculate the information I from its definition given in Eq. 4 in the main text.

S3. APPROXIMATE CHANNEL ACTIVITY

We estimate the expected activity $\langle a_n \rangle$ by the probability that the normalized excitations \hat{e}_n exceed the expected normalized threshold α . Since both the sensitivities S_{ni} and the normalized concentrations \hat{c}_i are approximately log-normally distributed, \hat{e}_n can also be approximated by a log-normal distribution (Fenton, 1960). The

associated probability distribution function reads

$$f(\hat{e}_n) = \frac{1}{\sqrt{2\pi} S_n \hat{e}_n} \exp \left[-\frac{(M_n - \ln(\hat{e}_n))^2}{2S_n^2} \right] \quad (\text{S10})$$

and the cumulative distribution function is

$$F(\hat{e}_n) = \frac{1}{2} \text{erfc} \left[\frac{M_n - \ln(\hat{e}_n)}{\sqrt{2} S_n} \right]. \quad (\text{S11})$$

The parameters M_n and S_n can be determined from the mean and variance

$$\langle \hat{e}_n \rangle = \exp \left(M_n + \frac{S_n^2}{2} \right) \quad (\text{S12a})$$

$$\text{var}(\hat{e}_n) = e^{2M_n + S_n^2} (e^{S_n^2} - 1). \quad (\text{S12b})$$

Solving these equations for M_n and S_n , we obtain

$$M_n = \ln \langle \hat{e}_n \rangle - \zeta \quad \text{and} \quad S_n = \sqrt{2\zeta}, \quad (\text{S13})$$

where $\zeta = \frac{1}{2} \ln(1 + \text{var}(\hat{e}_n) \langle \hat{e}_n \rangle^{-2})$. Eq. 6 of the main text follows from this and Eq. 5. For small $\langle a_n \rangle$ we have

$$\langle a_n \rangle \approx \frac{2\sqrt{\zeta/\pi}}{\ln(\alpha) + \zeta} \exp \left[-\frac{(\ln(\alpha) + \zeta)^2}{4\zeta} \right], \quad (\text{S14})$$

which follows from $\text{erfc}(x) \approx e^{-x^2}/(x\sqrt{\pi})$, valid for $x \gg 1$. For small ζ , we obtain the approximate scaling $\ln \langle a_n \rangle \sim -(\ln \alpha)^2/(4\zeta)$, where $\zeta \sim s^{-1}$ for $s \gg 1$.

S4. ODOR DISCRIMINABILITY

We quantify the discriminability of two odors by the Hamming distance d of their respective representations \mathbf{a} for several different cases:

a. Uncorrelated odors The expected distance $\langle d \rangle$ between the activity patterns $\mathbf{a}^{(1)}$ and $\mathbf{a}^{(2)}$ of two independent odors is

$$\langle d \rangle = N_R \left(\langle a_n^{(1)} \rangle + \langle a_n^{(2)} \rangle - 2\langle a_n^{(1)} \rangle \langle a_n^{(2)} \rangle \right), \quad (\text{S15})$$

where $\langle a_n^{(1)} \rangle$ and $\langle a_n^{(2)} \rangle$ denote the expected activities of the two odors, averaged over sensitivity matrices, and we neglect correlations $\text{cov}(a_n, a_m)$ for simplicity.

b. Adding target to background We calculate the expected change $\langle d \rangle$ of the representation when a target odor \mathbf{c}^t is added to a background odor \mathbf{c}^b . Because the odor concentrations are specified, we consider the actual excitations e_n instead of the normalized quantities \hat{e}_n . Taking an ensemble average over sensitivity matrices, the excitations associated with the two odors are characterized by probability distribution functions $f_E^t(e^t)$ and

$f_E^b(e^b)$ for the target and the background, respectively. We here consider log-normally distributed e_n , which are parameterized by their mean and variance,

$$\langle e_n \rangle = \bar{S} \sum_{i=1}^{N_L} c_i \quad \text{var}(e_n) = \text{var}(S_{ni}) \sum_{i=1}^{N_L} c_i^2, \quad (\text{S16})$$

where $\text{var}(S_{ni}) = \bar{S}^2(e^{\lambda^2} - 1)$.

When the target is added to the background, the expected threshold $\langle \gamma \rangle$ increases from $\gamma^b = \alpha \langle e^b \rangle$ to $\gamma^s = \alpha(\langle e^b \rangle + \langle e^t \rangle)$, where $\langle e^\kappa \rangle$ denotes the mean excitation $\langle e^\kappa \rangle = \int z f_E^\kappa(z) dz$ for $\kappa = t, b$. This increase in the threshold can deactivate a channel if it was previously active, i.e. if its excitation was larger than the threshold associated with the background, $e^b > \gamma^b$. For such e^b , the probability that the receptor gets deactivated by adding the target is $P(e^b + e^t < \gamma^s | e^b)$. Integrating over all possible e^b , we thus get the probability p_{off} that a channel becomes inactive,

$$\begin{aligned} p_{\text{off}} &= \int_{\gamma^b}^{\infty} P(e^b + e^t < \gamma^s | e^b) f_E^b(e^b) de^b \\ &= \int_{\gamma^b}^{\infty} F_E^t(\gamma^s - e^b) f_E^b(e^b) de^b, \end{aligned} \quad (\text{S17})$$

where $F_E^t(e^t)$ is the cumulative distribution function associated with $f_E^t(e^t)$. Conversely, a channel becomes active when the additional excitation by the target odor brings it above the threshold γ^s . The associated probability p_{on} reads

$$p_{\text{on}} = \int_0^{\gamma^b} [1 - F_E^t(\gamma^s - e^b)] f_E^b(e^b) de^b. \quad (\text{S18})$$

Taken together, the expected number $\langle d \rangle$ of channels that change their state reads

$$\langle d \rangle = N_R \cdot (p_{\text{on}} + p_{\text{off}}). \quad (\text{S19})$$

There are three simple limits that we can solve analytically: If there is no target, $\langle e^t \rangle = 0$, the activation pattern does not change and we have $\langle d \rangle = 0$. In the opposing limit of a dominant target, $\langle e^t \rangle \rightarrow \infty$, the activation patterns are independent and we recover the distance $\langle d \rangle_{\text{max}}$ for uncorrelated odors, which is given by Eq. S15. Lastly, in the case where the target and the background are identically distributed, $\langle e^b \rangle = \langle e^t \rangle$ and $\text{var}(e^b) = \text{var}(e^t)$, we have $\langle d \rangle = \frac{1}{2} \langle d \rangle_{\text{max}}$.

c. Discriminating two odors of equal size We consider the simple case of two odors that each contain s ligands at equal concentration, sharing s_b of them, such that the expected threshold $\langle \gamma \rangle$ is the same for both odors. Similar to the derivation above, we here calculate the probability p that a channel is active for one odor, but not for the other. The s_b ligands that are present in both odors

cause a baseline excitation e^b , which is distributed according to $f_E^b(e^b)$. A channel is inactive for an odor with probability $F_E^d(\langle \gamma \rangle - e^b)$, where $F_E^d(e^d)$ is the cumulative distribution function of the excitation caused by the $s_d = s - s_b$ different ligands. Hence,

$$p = 2 \int_0^{\langle \gamma \rangle} F_E^d(z) [1 - F_E^d(z)] f_E^b(e^b) de^b, \quad (\text{S20})$$

where $z = \langle \gamma \rangle - e^b$. Note that the upper bound of the integral is $\langle \gamma \rangle$ since channels will be active for both odors if $e^b \geq \langle \gamma \rangle$. The associated Hamming distance $\langle d \rangle$ between the two odors is then given by $\langle d \rangle = p N_R$.

S5. RECEPTOR BINDING MODEL

We consider a simple model where receptors R_n get activated when they bind ligands L_i . This binding is described by the chemical reaction $R_n + L_i \rightleftharpoons R_n L_i$, where $R_n L_i$ is the receptor-ligand complex. In equilibrium, the concentrations denoted by square brackets obey $[R_n L_i] = K_{ni} \cdot [R_n][L_i]$, where K_{ni} is the binding constant of the reaction. Hence,

$$[R_n L_i] = \frac{c_n^{\text{rec}} K_{ni} c_i}{1 + \sum_i K_{ni} c_i}, \quad (\text{S21})$$

where we consider the case where multiple ligands compete for the same receptor. Here, $c_i = [L_i]$ is the concentration of free ligands and $c_n^{\text{rec}} = [R_n] + \sum_i [R_n L_i]$ denotes the fixed concentration of receptors, which is related to the copy number of receptors of type n . We consider a simple receptor model where the excitation is proportional to the concentration of the bound ligands, such that the excitation accumulated in glomerulus n reads

$$e_n = \beta_n \frac{N_n^{\text{rec}}}{c_n^{\text{rec}}} \sum_{i=1}^{N_L} [R_n L_i] = \beta_n N_n^{\text{rec}} \frac{\sum_i K_{ni} c_i}{1 + \sum_i K_{ni} c_i}. \quad (\text{S22})$$

Here, N_n^{rec} is the copy number of receptors of type n and β_n characterizes their excitability, which could for instance be modified by point mutations (Yu *et al.*, 2015). Defining $S_{ni} = \beta_n N_n^{\text{rec}} K_{ni}$, we recover Eq. 1 of the main text in the limit of small concentrations, $\sum_i K_{ni} c_i \ll 1$. The sensitivities are thus proportional to the copy number N_n^{rec} and the biochemical details encoded in $\beta_n K_{ni}$.

REFERENCES

- Fenton, L. F., Communications Systems, IRE Transactions on **8**, 57 (1960).
- Wu, J., Mehta, N. B., and Zhang, J., in *GLOBECOM '05. IEEE Global Telecommunications Conference, 2005.*, Vol. 6 (2005) pp. 3413–3417.
- Yu, Y., Claire, A., Ni, M. J., Adipietro, K. A., Golebiowski, J., Matsunami, H., and Ma, M., Proc. Natl. Acad. Sci. USA **112**, 14966 (2015).